

Recommendations for the Future of Evaluation Within COSEE

Evaluators for the COSEE Centers are integral partners with the Principal Investigators (PIs) and staff in the creation and expansion of COSEE's mission—*engaging scientists and educators to transform ocean sciences education for all*.

Our evaluation efforts have focused, and will continue to focus, on identifying and measuring the effectiveness of COSEE programs, activities, and resources on ocean scientists as they develop proposals, implement actions for broader impacts, design and disseminate innovative educational materials/methods, and teach diverse audiences. Additionally, evaluation work will investigate and measure the impact on formal and informal educators as they integrate ocean literacy principles and concepts into the learning experiences they design and implement. This document, developed by the COSEE Evaluation Working Group (EWG), proposes a multi-layered approach for internal evaluation of COSEE programs and activities to meet the diverse needs of NSF, as well as COSEE scientists, educators, PIs and evaluators. This document does not address external evaluation of the COSEE Network.

Evaluation Task I. Refine and Expand Data Collection and Aggregation Infrastructure

A uniform set of surveys for scientist and educator engagement was a major accomplishment in 2009 and 2010. It is critical that timely, consistent, transparent, and accessible data regarding COSEE's scope, reach, and impact be available to support accountability, evaluation, and research needs. To date, efforts in this area have included the above surveys, an activity database, and meta-analysis of NSF and Center evaluation reports. Moving forward, we suggest a proactive and coordinated response.

- **Expanding and analyzing this rigorous, data-driven approach to impacts, program fidelity, and outcomes.**

Future data-collection efforts must be aimed at gathering consistent data regarding the scale and scope of impact (e.g., numbers of scientists served, modes of service, etc.) as well as the broader impact of programs and services.

- **Development and maintenance of cross-Center database systems that can support the efficient collection, aggregation, and disaggregation of these data. These systems will need to be implemented and supported across the COSEE Network.**

Online information systems powered by back-end databases are available that would support both accountability and administrative needs and could be tailored for the specific needs of the COSEE community. Information systems could be designed to collect information about scientists, educators, and program participants and track unduplicated participants nationally to better understand the scope and reach of COSEE. The cross-Center database would need to easily accept data in standardized formats used by Centers (avoiding re-entry of data) and output data in standardized formats. Expanded data collection and aggregation infrastructure would require additional resources beyond those currently possessed by Centers and the Network.

Evaluation Task II. Continue Center Evaluations for Program Fidelity & Improvement

Evaluation efforts designed for COSEE Centers have made up the majority of past evaluation efforts. These have focused on supporting program development and improvement through extensive needs assessment, formative, and summative evaluations. Data collection tools and methods are often tailored to particular programmatic foci. Results of these, mostly formative, evaluation activities support Center accountability efforts by providing rich, detailed accounts of program design, implementation, and impacts. These evaluations have informed PIs and other decision makers about the demonstrated needs of scientists and educators and the outcomes of programs and activities, enabling decision makers to make substantive changes to improve the quality of COSEE activities over the past nine years.

- **Continuation and augmentation of these efforts to refine the understanding of scientists' and graduate students' needs, plus measure program fidelity, outputs, performance, and impact on scientists, educators and the public.**
- **Refining the goals of COSEE and establishing related benchmarks as well as standards for measurement of progress and impact towards them.**

These evaluations would expand to support evidence-based program design and development, especially related to new or especially promising initiatives. Additional funding may be required at Centers to augment Center evaluation activities beyond their current levels.

Evaluation Task III. Improve and Expand Cross-Center Comparative Studies

With the increased budget allocated for Network evaluation activities in the most recent round of COSEE funding, there is an opportunity to clarify priorities and desired outcomes for cross-Center comparative studies. Exemplary programs, such as COS/COSIA and concept mapping, disseminated broadly across the Network, will benefit from rigorous measurement of identifiable components and outcomes that lead to models for effective practices for COSEE and more broadly. Based on the COSEE Network, these studies could use experimental and/or quasi-experimental designs to assess the effectiveness of different formats on different audiences.

Increasing COSEE's capacity to scale up programs or activities, and assess their impact on ocean sciences education or research, would be one possible area of cross-Center evaluation.

- **Funding the EWG or a similar body, as well as individual COSEE evaluators to initiate, refine, and implement rigorous, exemplary cross-Center evaluation efforts.**
- **Continuing and expanding investigations into the differentiated efficacy of COSEE initiatives and programs with various regional communities, as well as sub-groups such as graduate students, late career scientists, underrepresented groups, etc.**

Evaluation Task IV. Evaluation Research to Answer Efficacy and Effectiveness Questions

The EWG believes that COSEE could embark on evaluation research designed to provide rigorous evidence of programmatic impact. An emphasis on evaluation research studies could be

prioritized to meet more rigorous evaluation and data collection standards. COSEE does not currently have funding for evaluation research.

Utilization of cutting-edge research methods and understanding about learning could be added to research efforts to measure program impact.

- **Addition of rigorous research efforts, especially related to a limited number of more mature programs or initiatives that have already been a part of extensive COSEE formative evaluation efforts.**

For example, from COSEE CA the *Communicating Ocean Sciences* courses and the *Ocean Literacy Scope and Sequence* curricular materials have been very successfully and rigorously evaluated. These evaluations could support development of tool kits to effectively disseminate quality resources, such as best practices for teacher professional development or developing scientists' broader impact proposals. An *Ocean Literacy Index* (currently in development by COSEE CA) could also better measure the degree of implementation of the *Ocean Literacy Scope and Sequence* curricula in K-12 classrooms and informal education centers.

- **COSEE community should define and delineate research questions related to efficacy and effectiveness that warrant further investigation.**

Research of this kind extends beyond previous evaluation efforts, yet is a natural continuation of the increasingly sophisticated and collaborative evaluation efforts underway for COSEE Centers. Evaluators recognize the need for rigorous research to support the measurement of program outcomes and impacts, and will endeavor to continue to meet the scientific community's needs for clear measurements of those. The EWG and Center evaluators can periodically review or revise the proposed cross-Center evaluation plan to ensure that the most pressing research questions of the ocean sciences research and education communities are sufficiently addressed.

V. Deliverables and Expectations

A. Center Evaluation (up to 5% of program budget, \leq 50% of evaluators' time)

1. Formative evaluation as necessary for emerging programs and Centers
2. Summative evaluation to analyze impact of Center programs and activities on scientists, the ocean science research community, and other specified audiences
3. Input into annual report (possible significant overlap with 1 and 2)
4. Data collection would be tied to existing engagement survey categories or instruments as much as possible
5. With additional funding:
 - Initiate and refine research evaluation efforts as defined above
 - Publish evaluation findings in peer-reviewed articles

B. National COSEE and Cross-Center Evaluation (5% of program budget, \geq 50% of evaluators' time)

1. Biennial scientist engagement survey with analysis (beginning the cycle after completion of 2011 survey in January) administered by each Center, and maintained within the COSEE Network database.
2. Biennial educator engagement survey with analysis administered by each Center, again maintained within the COSEE Network database.
3. The engagement surveys would be coordinated, analyzed and summarized by a *funded* EWG. The current survey database is managed by the EWG and be incorporated into the COSEE Network database.
4. Evaluation of cross-Center projects and activities, including efficacy and effectiveness studies.
5. Number of times COSEE is mentioned in NSF ocean sciences research proposals [this has been done in the past by NSF]
6. Definition and implementation of goals, benchmarks, and related standards for measurement of *progress* and *impact*. These standards would be developed by the EWG and agreed upon by the COSEE Council.
7. Annual evaluators meeting (≤ 2 days)
8. With additional funding:
 - Annual evaluation research progress report (formative evaluation) with coordination and support from EWG
 - Annual supplemental (reduced in size) scientist engagement data collection (survey)
 - Development of on-line data collection tools, integrated into centralized database, to facilitate accurate and timely data collection from scientists, educators and other participants nationwide.

C. Evaluation Goals

1. Longitudinal data collection and analysis on scientist and educator engagement in COSEE with impact on scientists, educators, the public, and other key audiences.
2. A centralized survey information system for data collection, hosting, management and analysis done by the EWG or selected COSEE Center evaluators.
3. Consensus acceptance by COSEE Centers of standards of measurement for key metrics including impact on scientists, impact on educators and students, progress toward achieving complex goals, and changes in understanding of selected scientific concepts (including climate change.)
4. Publication of three peer-reviewed journal articles related to COSEE evaluation or evaluation research studies. Coordination, collaboration, and other support would be one of the responsibilities of the EWG once the decadal review is finished.

VI. Additional Recommendations

A. Evaluation Management

1. Expand and formalize the EWG to become an Evaluation Coordinating Group (ECG) with separate, sustained funding.
2. This ECG is envisioned to have responsibility for supporting collaborative evaluation efforts, possibly including evaluation research. Our recommended structure would be to include all evaluators in the ECG, to have a leadership team chosen from among them to coordinate activities and communication, with the addition of one staff person to provide office support. Additionally, fund an internal evaluator position to coordinate cross-Center evaluation efforts. This would be a funded position, simultaneously chairing the EWG/ECG, with length of service to be decided (rotating or permanent). The ECG could possibly be housed at one Center's partner organization or NCO.
3. Continue regular (biweekly) EWG/ECG conference calls.
4. Continue annual meeting of COSEE evaluators and monthly conference calls.
5. Include evaluation representative(s) at national COSEE meetings and Council calls.

B. Evaluation Resources

1. Reconsider funding priorities and standards for evaluation efforts to ensure adequate compensation, including travel, for tasks assigned to evaluators.
2. Funding for all Centers should reflect NSF's requirement that 5% of the total budget be allocated towards cross-Center and Network evaluation activities.
3. The EWG recommends that the remainder of Center evaluator time, equal to minimally an additional 5% of the total budget also be allocated towards Center evaluation activities to provide the necessary support to Center PIs and data for cross-Center and Network evaluation activities.
4. Should additional resources be available, we recommend funding for Center evaluation should be increased to at least 7%, and up to 10%, of Center budget to support enhanced collaboration, evaluation research, and publication.
5. Provide professional development for evaluators to accelerate implementation of increasingly refined and rigorous evaluation and research protocols, including experimental design and data analysis.

